Introduction to the course

Prerequirements

- OOP + Java
- Networks
- Database
- Basic knowedge of HTML/HTTP/CSS – only a quick catch up will be given
- Basic knowedge of Web servers (at least cookies, sessions, servlets – only a quick catch up will be given)

We’ll use Piazza

Students have to enrol using their unitn e-mail address at the following link:

https://piazza.com/unitn.it/fall2020/webarchitectures

Exam

Final project +

oral discussion on theory

or, as recommended:

Smaller projects during the course +

oral discussion on theory

START

Introduction to the Web Architectures course

Q

Why is Web project development different?

Web development challenges

- The functionality of a web application is distributed, with part running on the server near the application's data and part running in the client(browser) near the application users.

- Web applications must integrate a complex collection of diverse technologies, (HTML, JavaScript, HTTP, SSL and one or more server side languages such as PHP, Java, Python or Ruby). Typically different languages are used for server and client side programming.

- Web Applications must support variety of browsers with feature sets that are not still 100% compatible.

Web development challenges

- Web applications support multiple users with (what appears to be) a single server. Developers have to deal with concurrency and scalability. Popular web applications must handle 100 – 1000× the workload of traditional applications.

- Web applications must be highly customizable. In the pre-Web GUI world a uniform look and feel was encouraged, but web applications strive for unique appearance and behaviors. This makes it more difficult to create reusable components.

- The public accessibility of web applications introduces a variety of security and privacy issues. It is easy for unaware developers to create security loopholes such as SQL injection attacks.

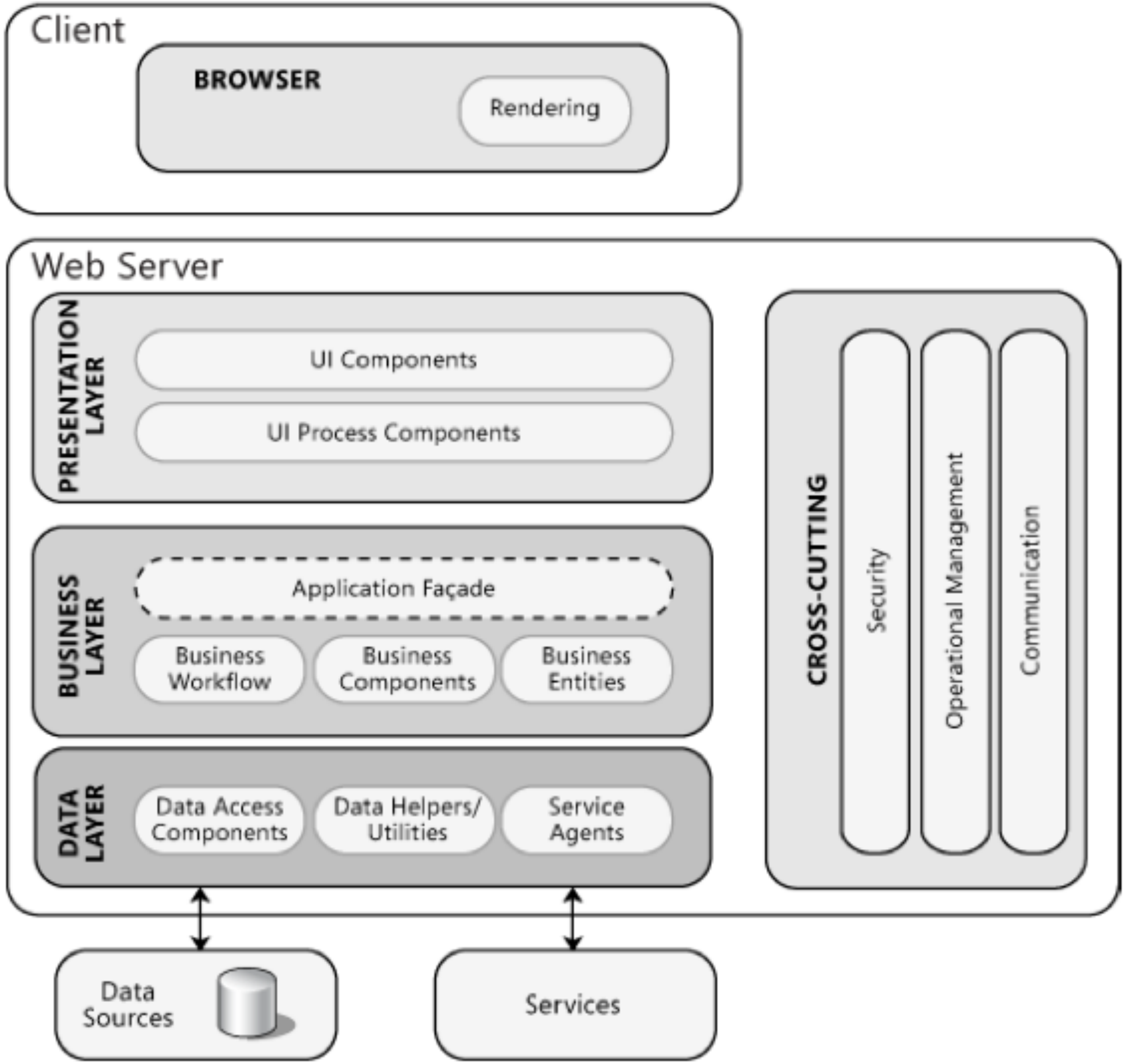
Q

What are the steps on a Web project development ?

Web development

- Model the domain.
- Define the Database, and database queries to persist and fetch data from the database
    - Create server-side code for accessing the database
    - Implement server-side code for any business logic
    - Implement an API that can be used to provide data for the presentation over an HTTP connection
    - Write integration tests for the API
- Start writing frontend code in JavaScript or similar technology
- Write client-side services to fetch data from the backend API
- Write client-side components to display the data on a web page.
- Build the page and style it
- Write some automated end-to-end tests for the web page.

Logical Web Architecture



Data Layer

abstracts the logic required to access the data source.

- Using a separate data layer makes the application easier to configure, maintain and hides the details of the data sources from other layers of the application.
- Entity objects can be populated by this layer or can be used to update the data sources.
- Data Transfer Objects(DTOs) are used to pass the data between layers.

Business Layer

implements the business logic of the application and long-running work-flows.

- A separate business layer:
  - improves the maintainability and testability of an application
  - allows to centralize and reuse common business logic functions.
- Business entities that represent the real world data are used to pass data between components.

Presentation layer

- **displays the user interface(UI)**
- **facilitates the user interaction**

user interaction logic is decoupled from the UI components.

The presentation layer consists of

- server-side components: prepares the HTML
- client side component: executes scripts and displays the HTML sent by the server side component.

Cross-cutting components

- **Security** for implementing functionality related with authentication, authorization, and validation.
- Operational Management for **exception handling** policies, **logging**, **performance counters**, **configuration**, and **tracing** tasks.
- Communication to **facilitate the communication with other services and applications**.

Q

What is the difference between the Web and the Internet?

What is the Internet?

- "The Internet is a global system of interconnected computer networks that interchange data by packet switching using the standardized Internet Protocol Suite (TCP/IP)."

- Thus, the Internet is a network of networks, defined by the TPC/IP standards (such as FTP, Telnet, SMTP…).

What is the web?

The Wold Wide Web is an **information system** where documents and other web resources are identified by **Uniform Resource Locators** (URLs, such as https://www.example.com/), which may be interlinked by **hypertext**, and are accessible over the Internet.

The resources of the WWW are

- described via the **Hypertext Markup Language** (HTML)
- transferred via the **Hypertext Transfer Protocol** (HTTP)
- accessed by users by a software app called a web browser
- published by a software app called a web server.

History of the web

- 1989-1990 – **Tim Berners-Lee** invents World Wide Web at CERN.

- On 30 April 1993, CERN put the World Wide Web software in the public domain. Later, CERN made a release available with an open license, a more sure way to maximize its dissemination.
- TB-L moved from CERN to the Massachusetts Institute of Technology in 1994 to found the **World Wide Web Consortium (W3C)**, an international community devoted to developing open web standards.

Q

What is a Server ?

Client and Server

from "Computer Networking: A Top-Down Approach", Kurose-Ross:

- The computers and other devices connected to the Internet are often referred to as end systems (**hosts**).
- Hosts are sometimes further divided into two categories: **clients** and **servers**.

- Informally, clients tend to be desktop and mobile PCs, smartphones, and so on, whereas servers tend to be more powerful machines that store and distribute Web pages, stream video, relay e-mail, and so on.

BAD DEFINITION

Servers ?

Server = Waiter, Client = Customer

Client and Server

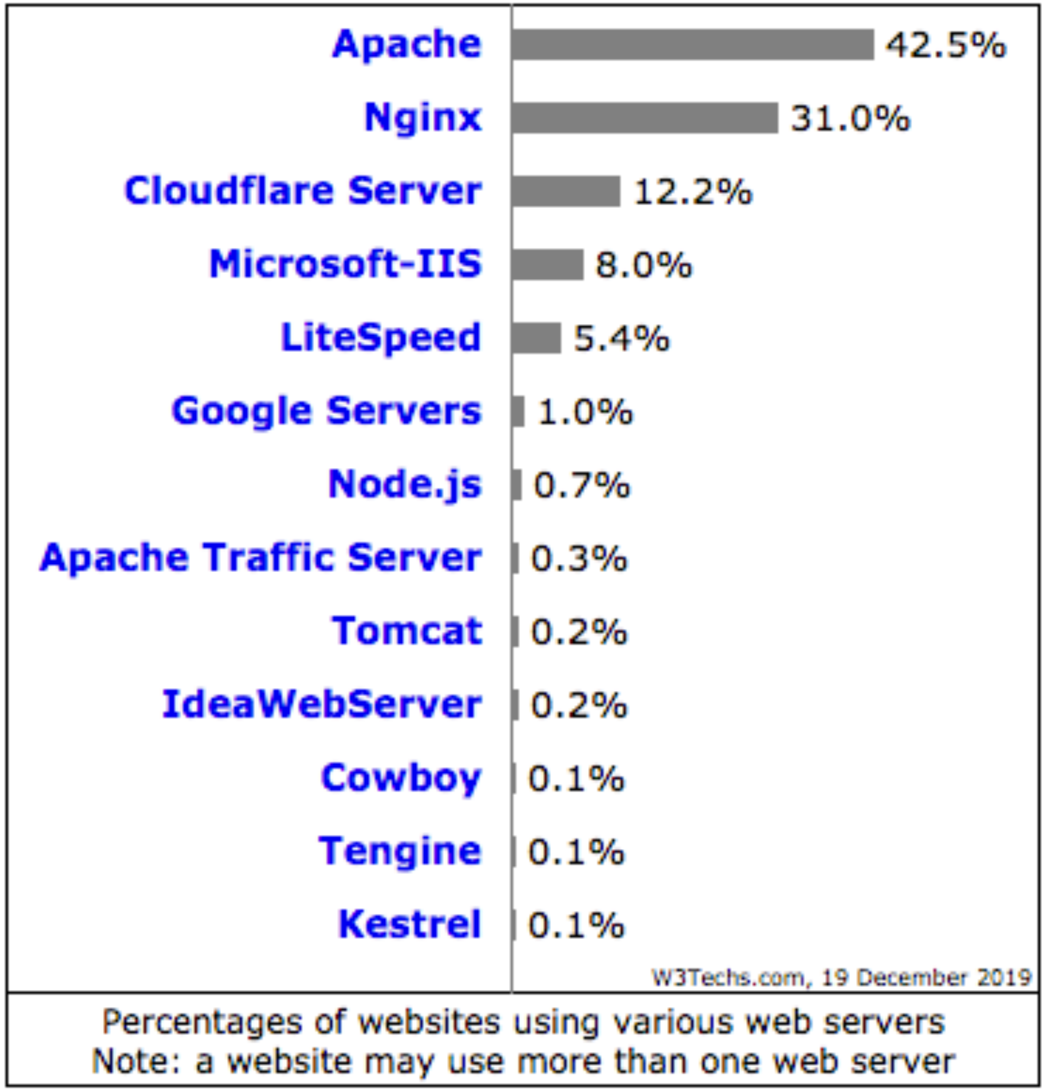
A better definition:

- A Server is a machine that opens a SocketServer connection, and waits for incoming calls (in order to provide a service)

- A Client is a machine the starts a connection (opening a Socket to the server) and requests a service.

Server and client are SOFTWARE roles, not HARDWARE concept!

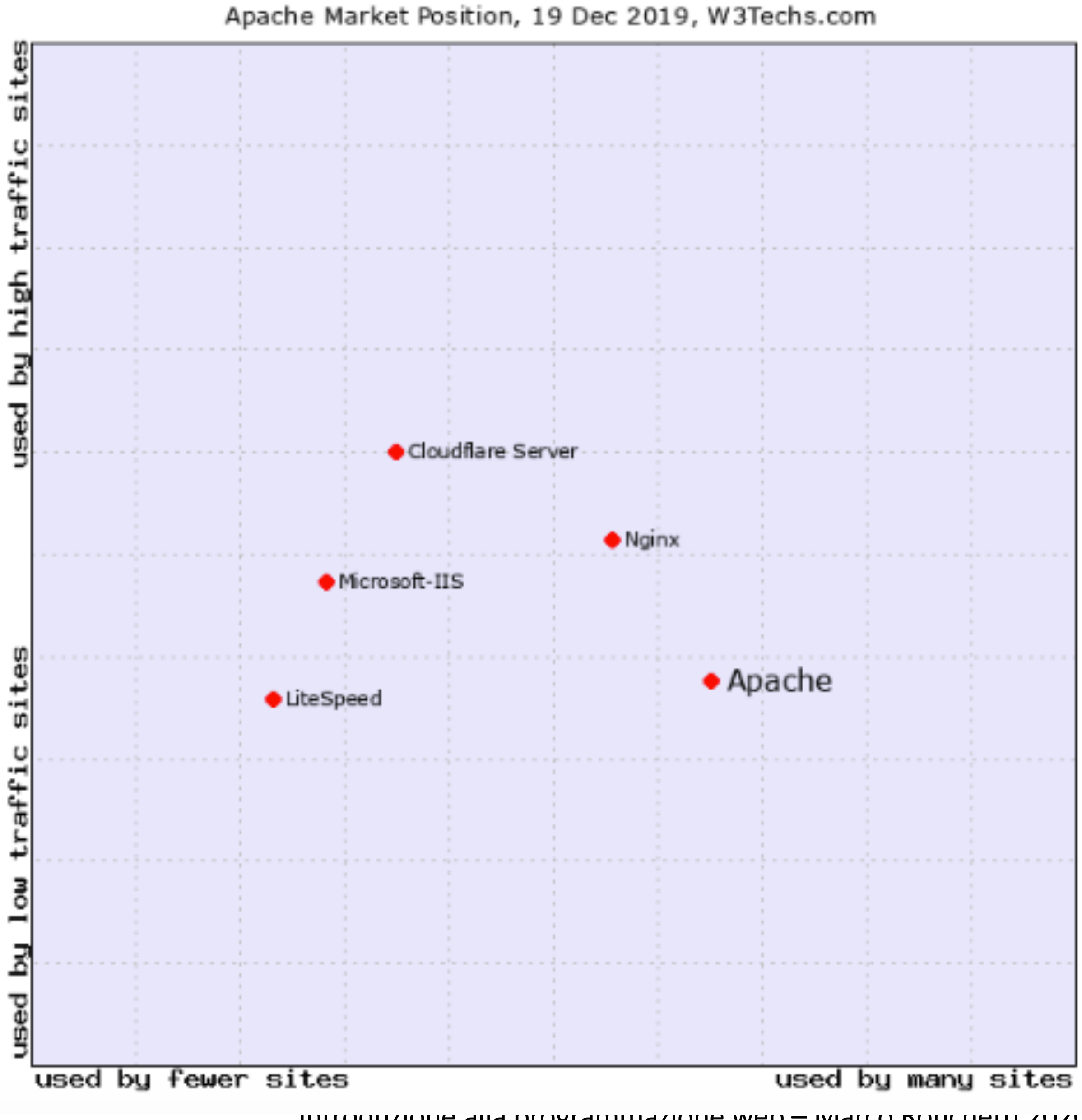
Servers on the web



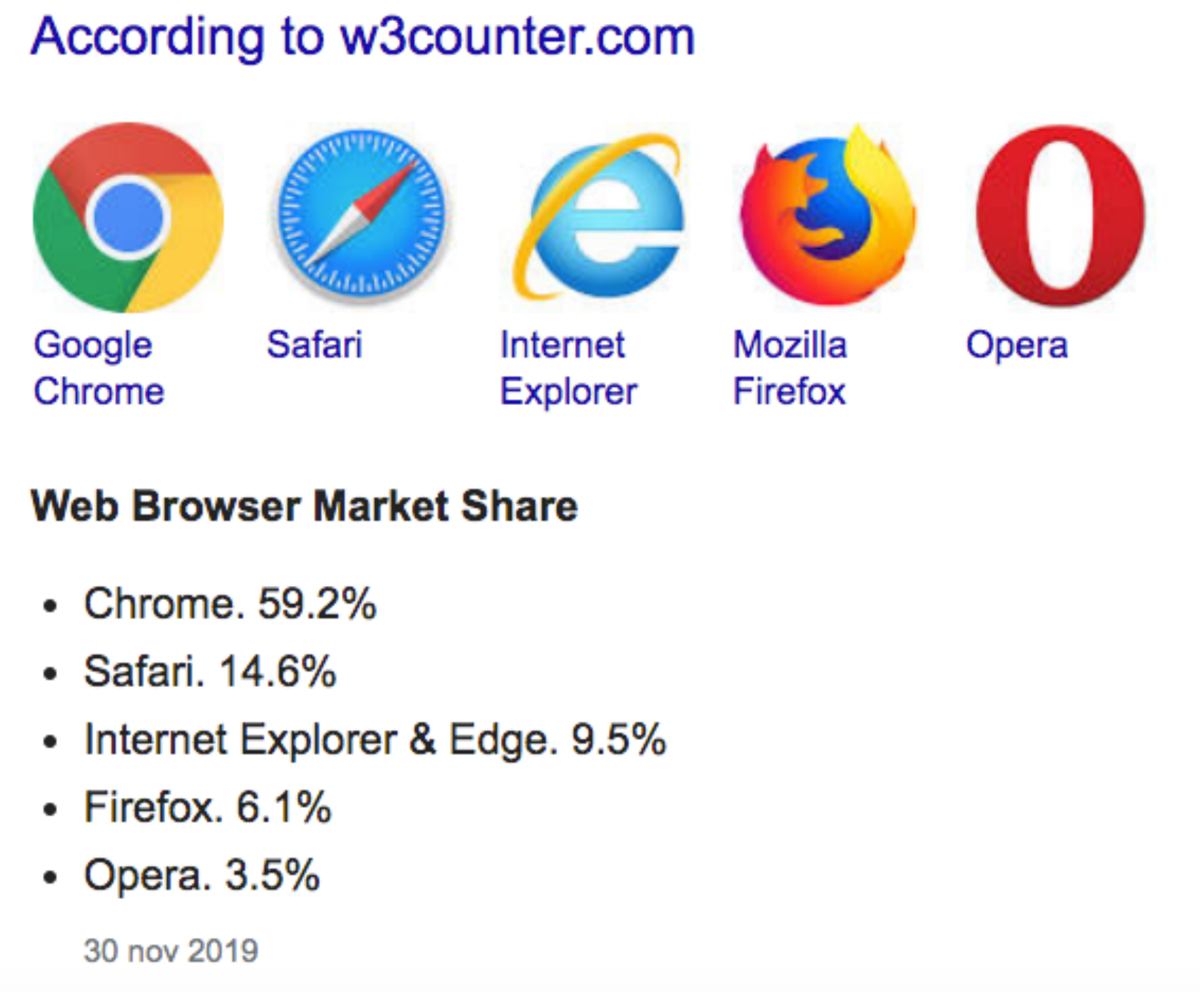
W3Techs.com, 19 December 2019

Percentages of websites using various web servers
Note: a website may use more than one web server

Servers on the web



Clients on the Web: Browsers



According to w3counter.com

Google Chrome · Safari · Internet Explorer · Mozilla Firefox · Opera

Web Browser Market Share

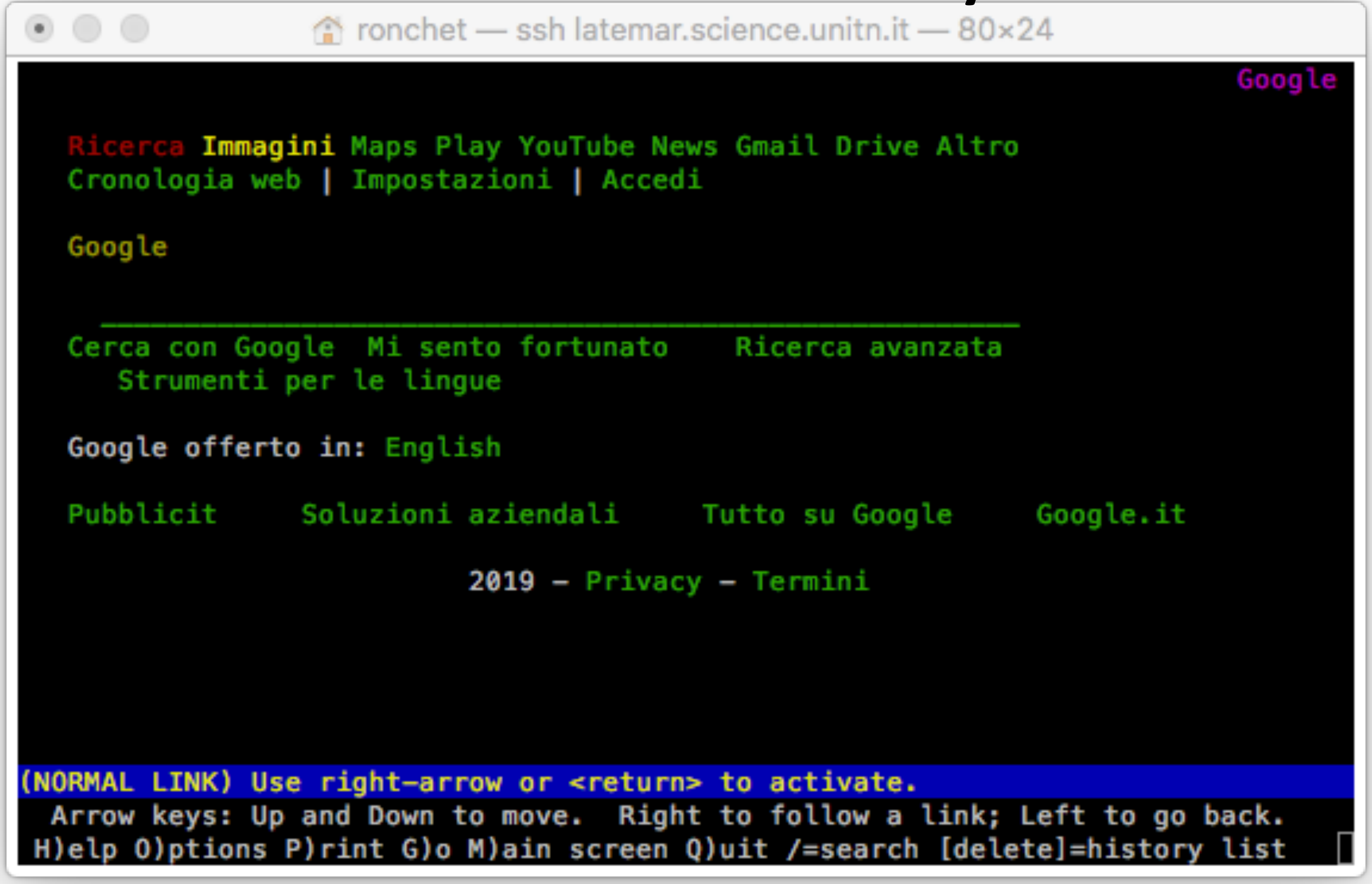
- Chrome. 59.2%
- Safari. 14.6%
- Internet Explorer & Edge. 9.5%
- Firefox. 6.1%
- Opera. 3.5%

30 nov 2019

Clients on the Web: Browsers

Could also be something else!

Lynx – text browser

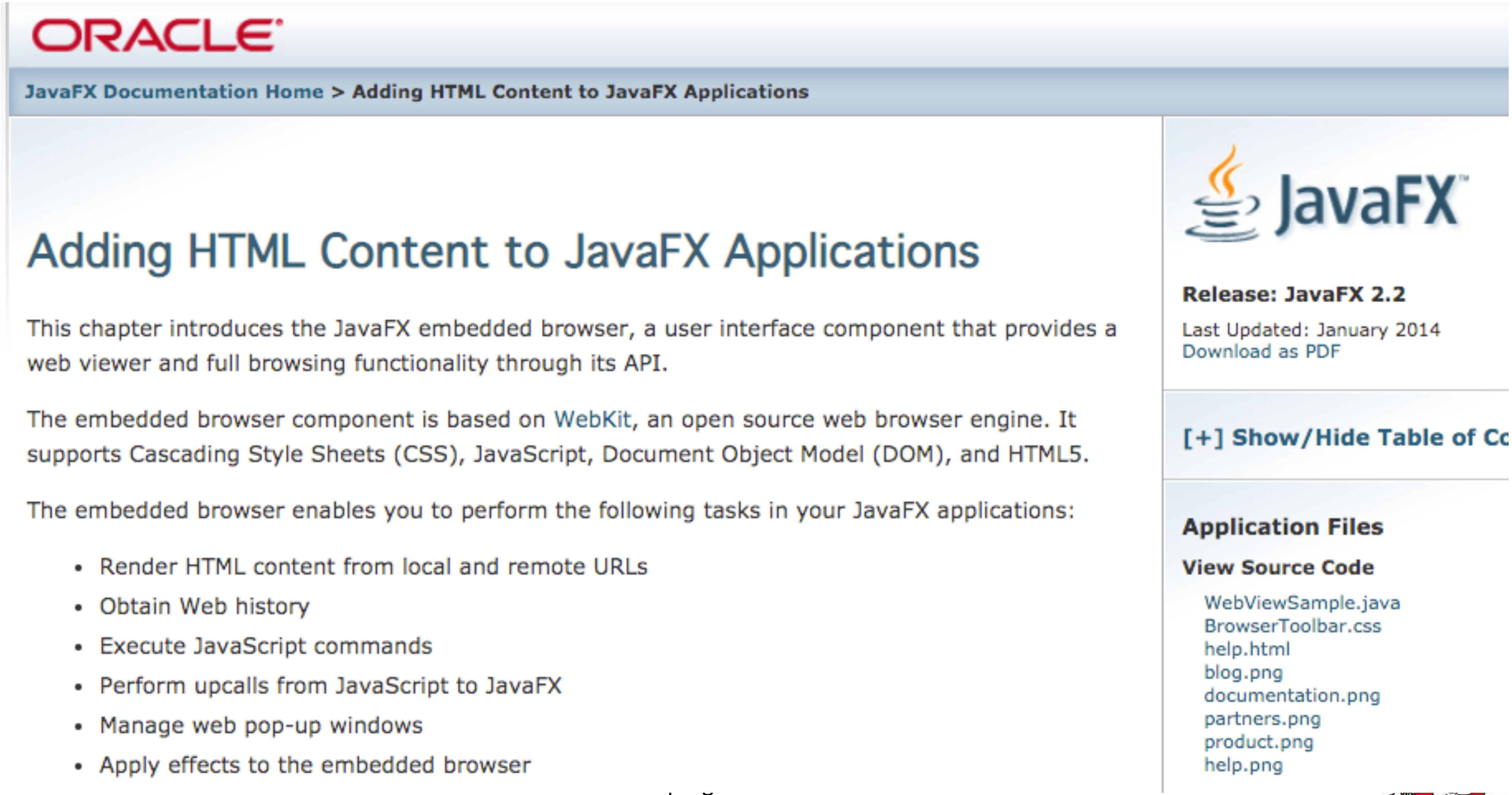


```
                                                                        Google


   Ricerca Immagini Maps Play YouTube News Gmail Drive Altro
   Cronologia web | Impostazioni | Accedi

   Google


     _______________________________________________________
   Cerca con Google  Mi sento fortunato    Ricerca avanzata
      Strumenti per le lingue

   Google offerto in: English

   Pubblicit     Soluzioni aziendali     Tutto su Google     Google.it

                           2019 - Privacy - Termini






(NORMAL LINK) Use right-arrow or <return> to activate.
  Arrow keys: Up and Down to move.  Right to follow a link; Left to go back.
 H)elp O)ptions P)rint G)o M)ain screen Q)uit /=search [delete]=history list
```

Clients on the Web: Browsers

Could also be something else!

WebView in JavaFX



ORACLE

JavaFX Documentation Home > **Adding HTML Content to JavaFX Applications**

Adding HTML Content to JavaFX Applications

This chapter introduces the JavaFX embedded browser, a user interface component that provides a web viewer and full browsing functionality through its API.

The embedded browser component is based on WebKit, an open source web browser engine. It supports Cascading Style Sheets (CSS), JavaScript, Document Object Model (DOM), and HTML5.

The embedded browser enables you to perform the following tasks in your JavaFX applications:

- Render HTML content from local and remote URLs
- Obtain Web history
- Execute JavaScript commands
- Perform upcalls from JavaScript to JavaFX
- Manage web pop-up windows
- Apply effects to the embedded browser

JavaFX

Release: JavaFX 2.2

Last Updated: January 2014
Download as PDF

[+] Show/Hide Table of Co

Application Files

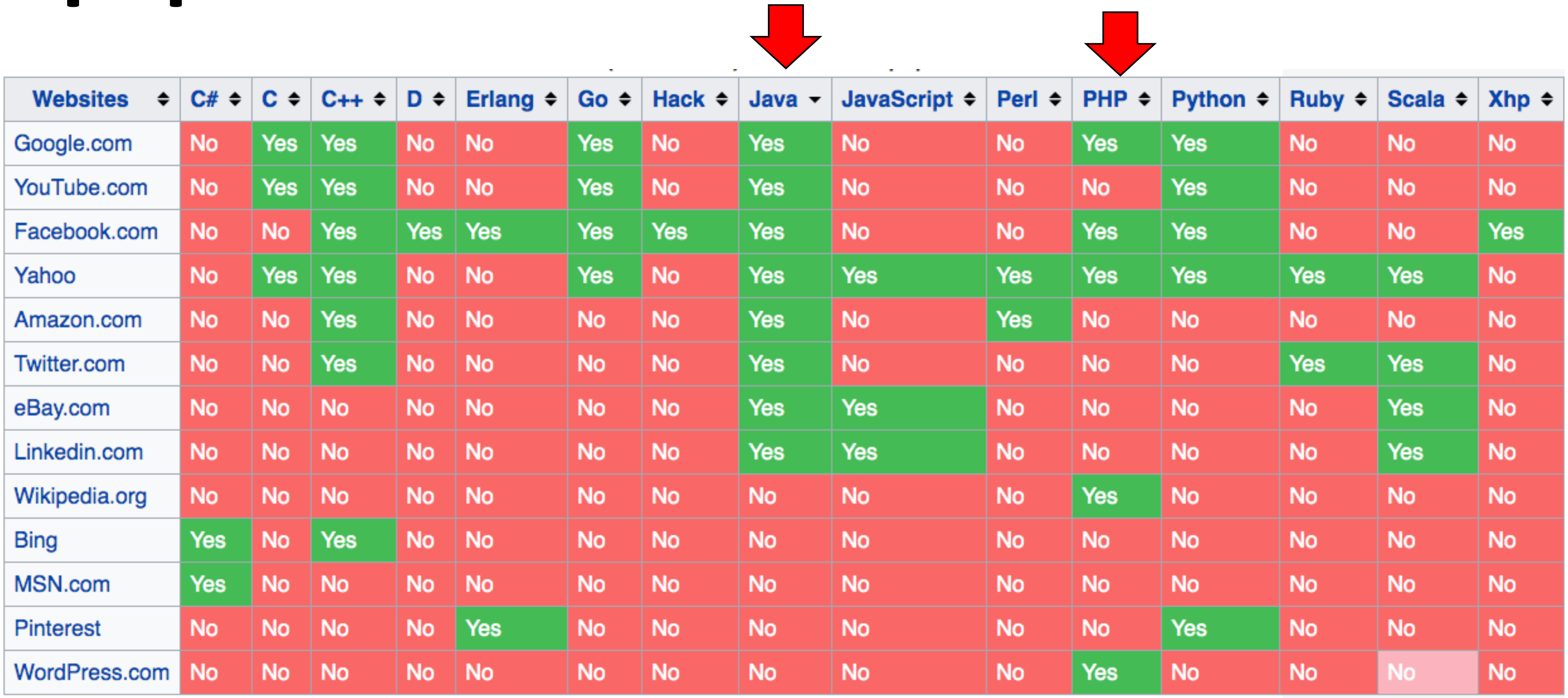
View Source Code

WebViewSample.java
BrowserToolbar.css
help.html
blog.png
documentation.png
partners.png
product.png
help.png

Q

Which languages are most used on the web ?

Back-end (Server-side) table in most popular websites



Websites	C#	C	C++	D	Erlang	Go	Hack	Java	JavaScript	Perl	PHP	Python	Ruby	Scala	Xhp
Google.com	No	Yes	Yes	No	No	Yes	No	Yes	No	No	Yes	Yes	No	No	No
YouTube.com	No	Yes	Yes	No	No	Yes	No	Yes	No	No	No	Yes	No	No	No
Facebook.com	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Yahoo	No	Yes	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Amazon.com	No	No	Yes	No	No	No	No	Yes	No	Yes	No	No	No	No	No
Twitter.com	No	No	Yes	No	No	No	No	Yes	No	No	No	No	Yes	Yes	No
eBay.com	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	Yes	No
Linkedin.com	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	Yes	No
Wikipedia.org	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No
Bing	Yes	No	Yes	No	No	No	No	No	No	No	No	No	No	No	No
MSN.com	Yes	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Pinterest	No	No	No	No	Yes	No	No	No	No	No	No	Yes	No	No	No
WordPress.com	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No
			7					8			5	5			

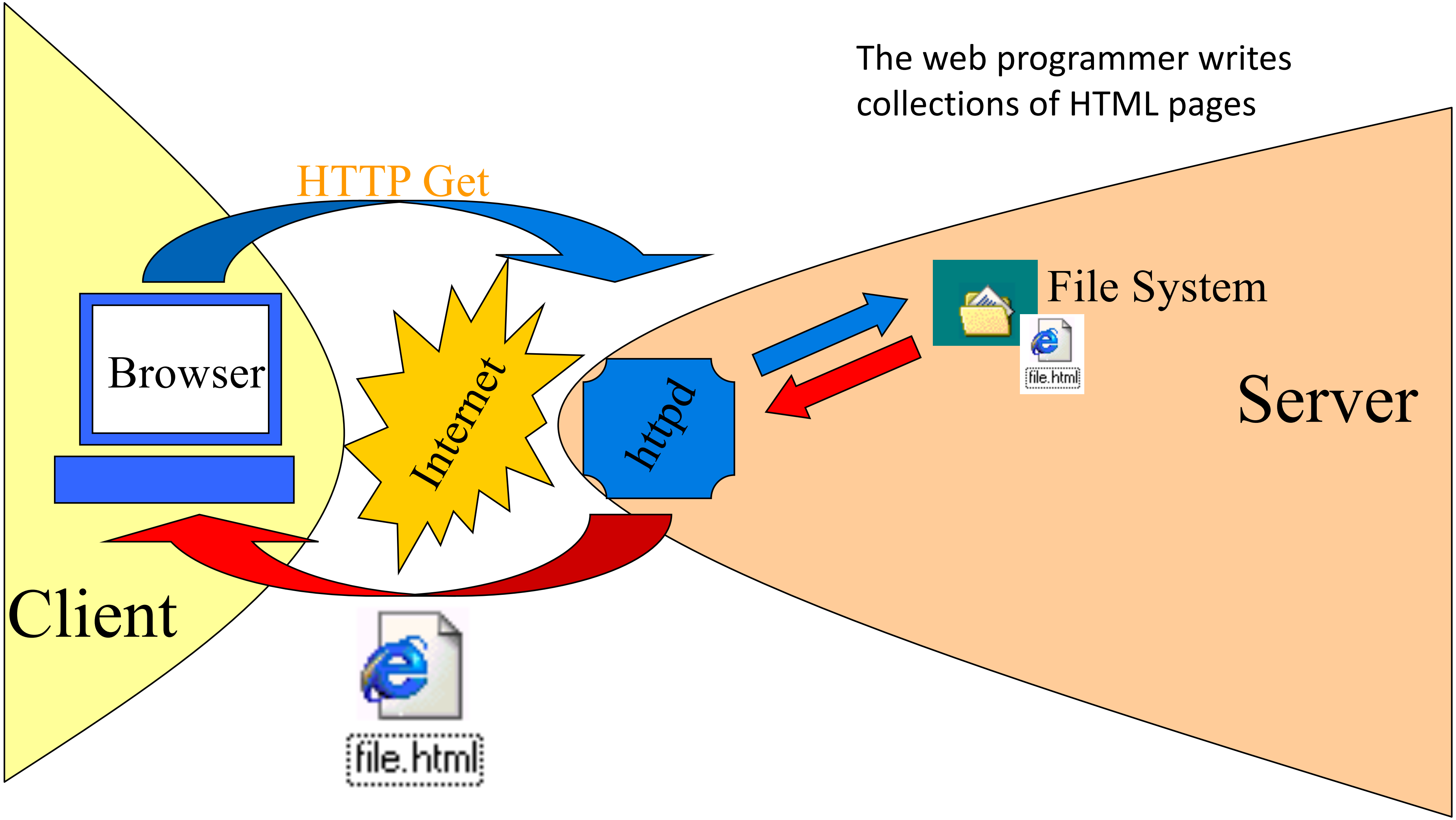
Front-end (client-side) table in most popular websites

Websites	Popularity (unique visitors per month)	Front-end	Notes
Google.com	1,600,000,000	JavaScript, TypeScript	The most used search engine in the world
Facebook.com	1,100,000,000	JavaScript	The most visited social networking site
YouTube.com	1,100,000,000	JavaScript	The most visited video sharing site
Yahoo	750,000,000	JavaScript	
Amazon.com	500,000,000	JavaScript	Popular internet shopping site
Wikipedia.org	475,000,000	JavaScript	"MediaWiki" is programmed in PHP; free online encyclopedia
Twitter.com	290,000,000	JavaScript	Popular social network.
Bing	285,000,000	JavaScript	Search engine from Microsoft.
eBay.com	285,000,000	JavaScript	Online auction house.
MSN.com	280,000,000	JavaScript	An email client, for simple use. Previously known as "messenger", not to be confused with Facebook's messaging platform.
Linkedin.com	260,000,000	JavaScript	World's largest professional network.
Pinterest	250,000,000	JavaScript	Search engine for ideas.
WordPress.com	240,000,000	JavaScript	Website manager software.

Q

What is the historical evolution of the web paradigm ?

The original web architecture: static pages

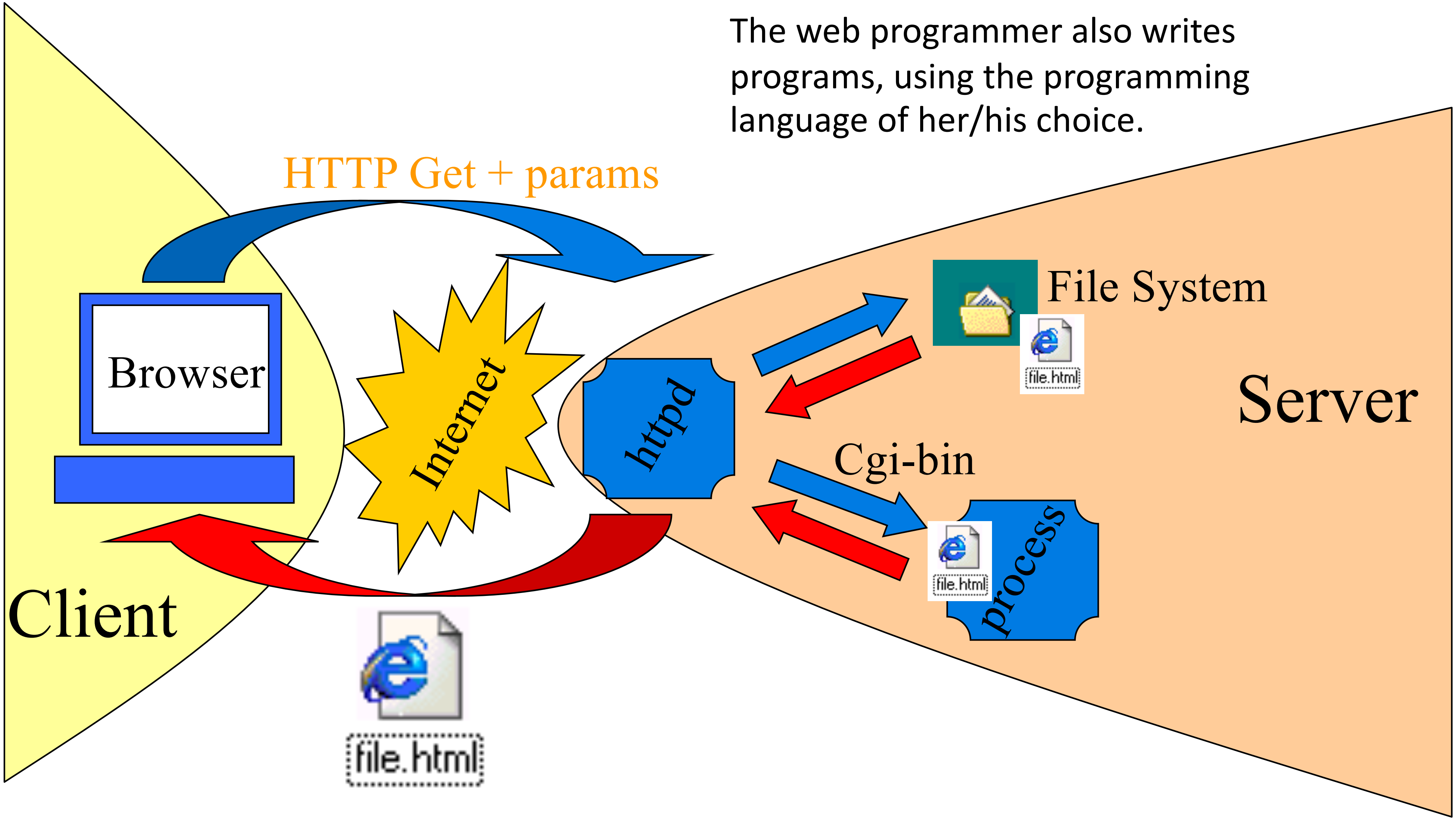


The web programmer writes collections of HTML pages

Initial idea: get (static) interlinked documents

The original web architecture: dynamic pages

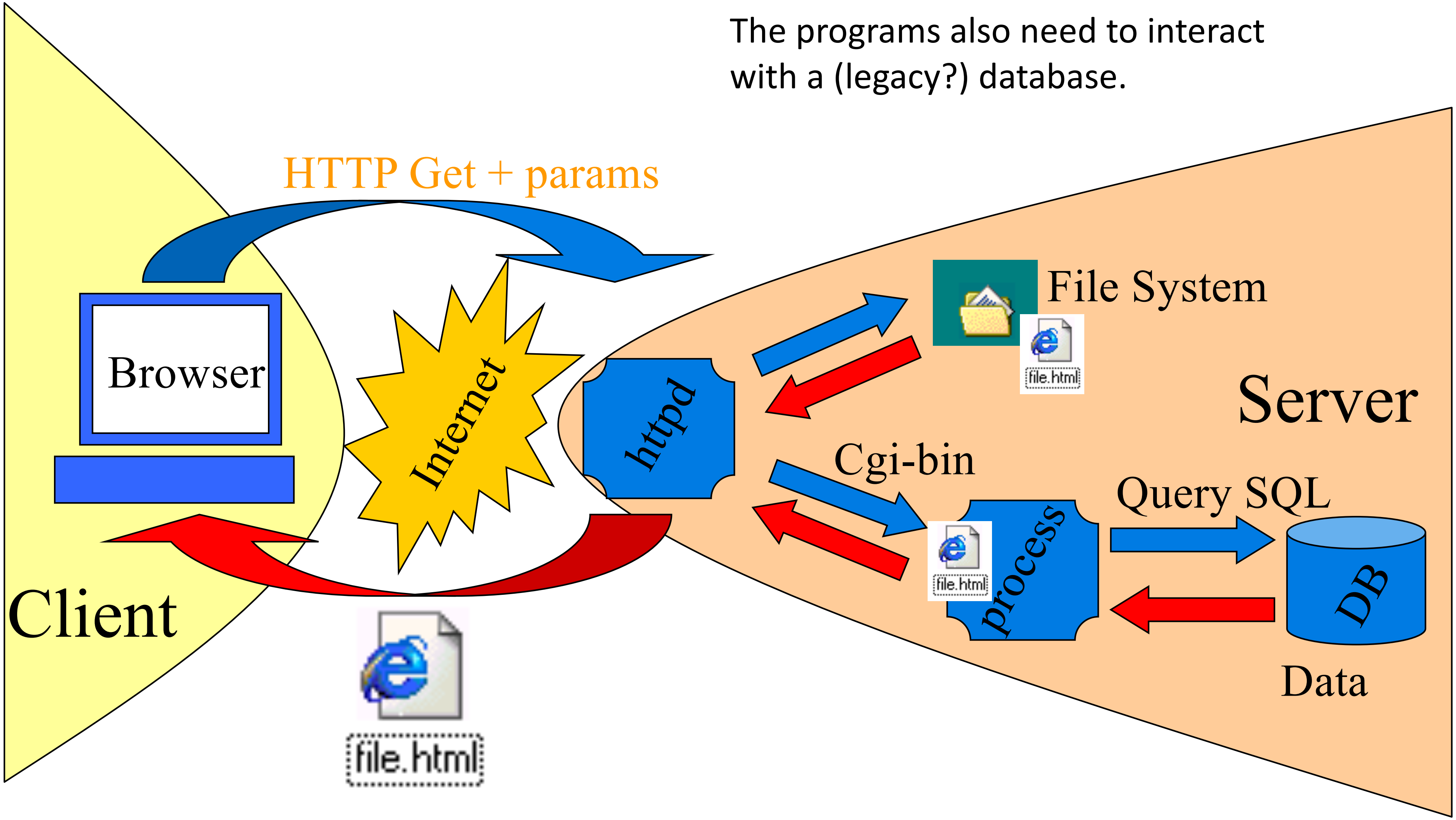
The web programmer also writes programs, using the programming language of her/his choice.



Evolution 1: dynamically create (interlinked) documents

The original web architecture: dynamic pages with DB

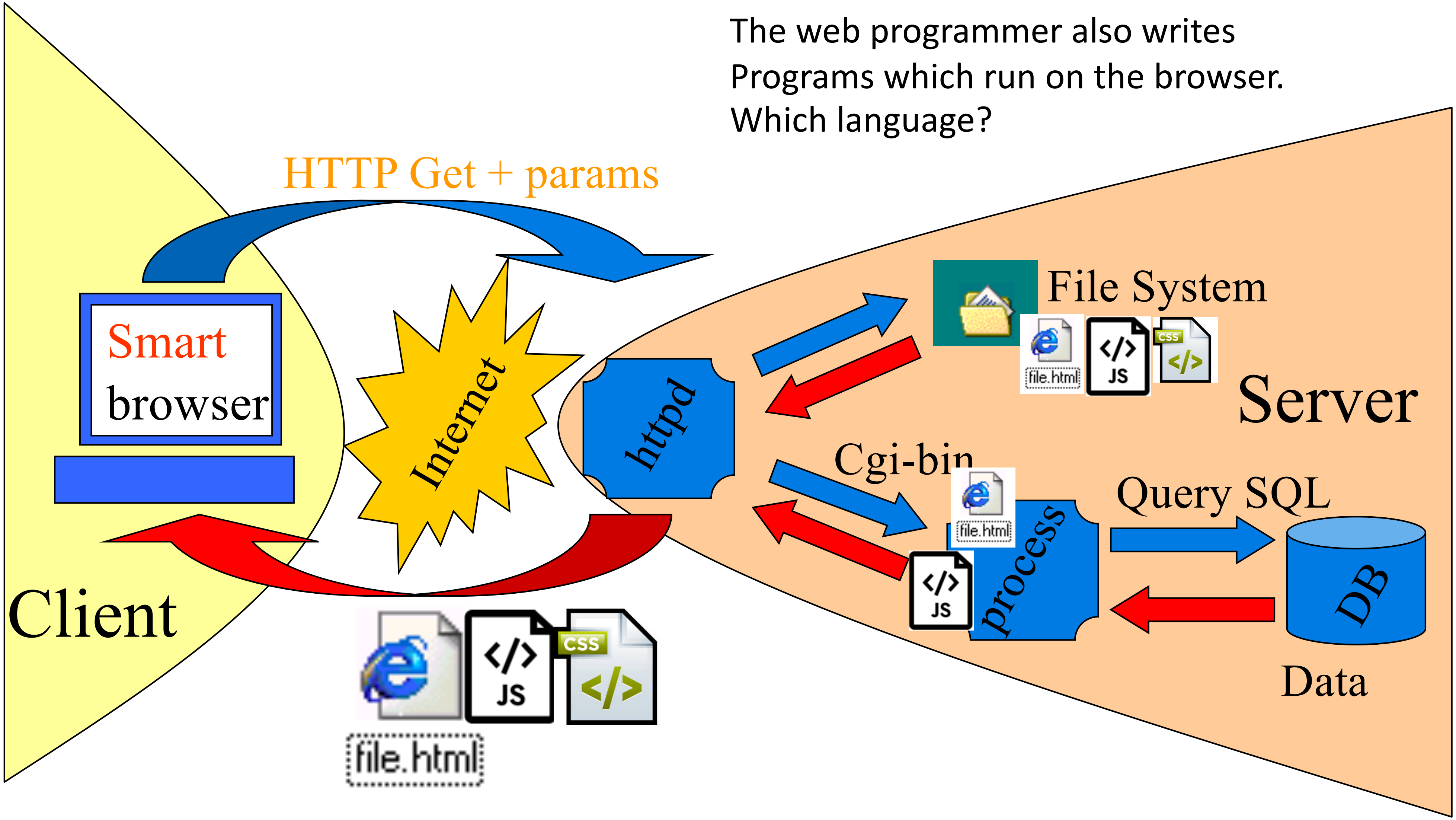
The programs also need to interact with a (legacy?) database.



Evolution 2: dynamically create (interlinked) documents interacting with a persistent data storage

The web architecture with smart browser

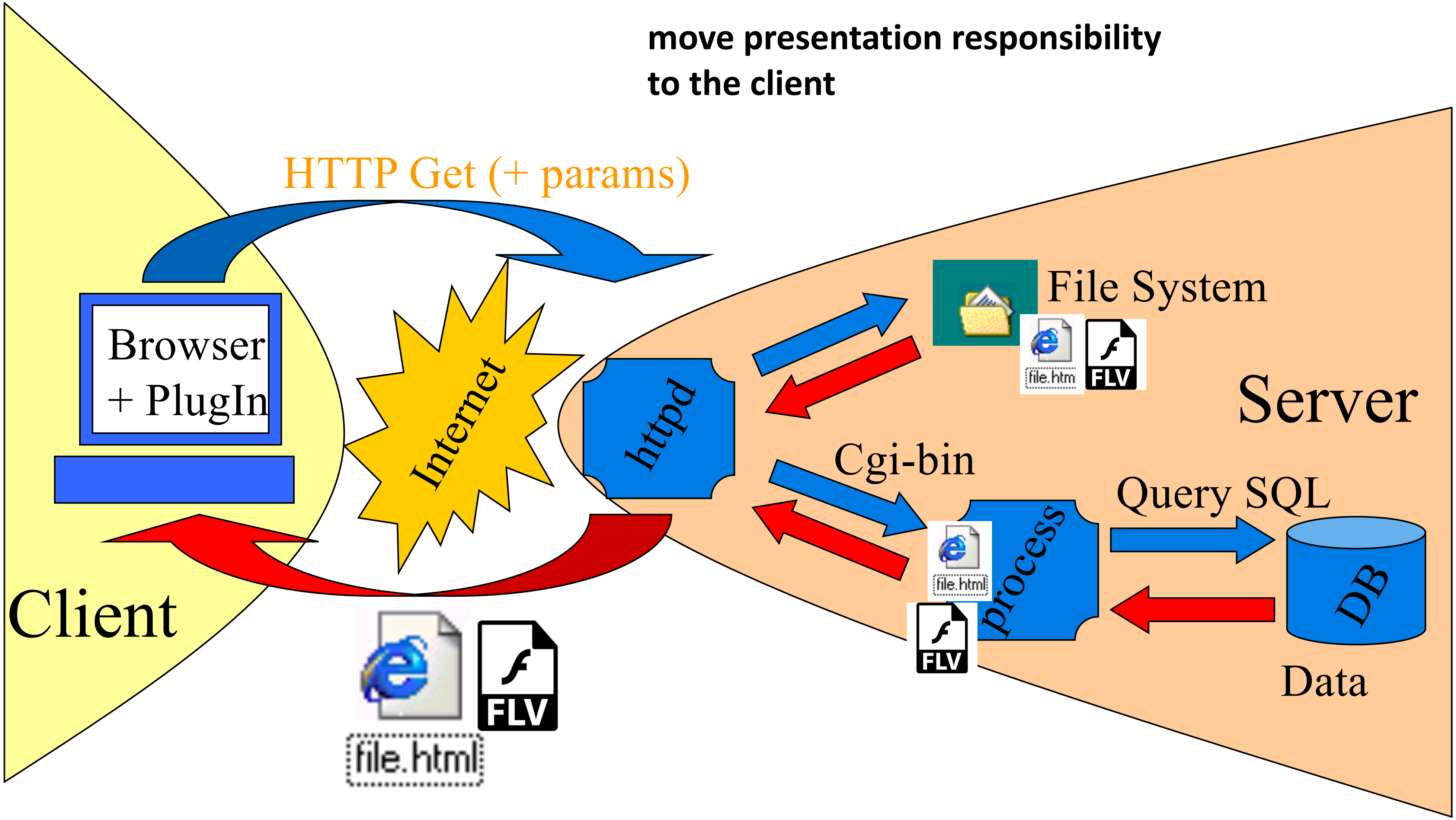
The web programmer also writes Programs which run on the browser. Which language?



Evolution 3: execute code also on client! (How ?)

The web architecture: plug in

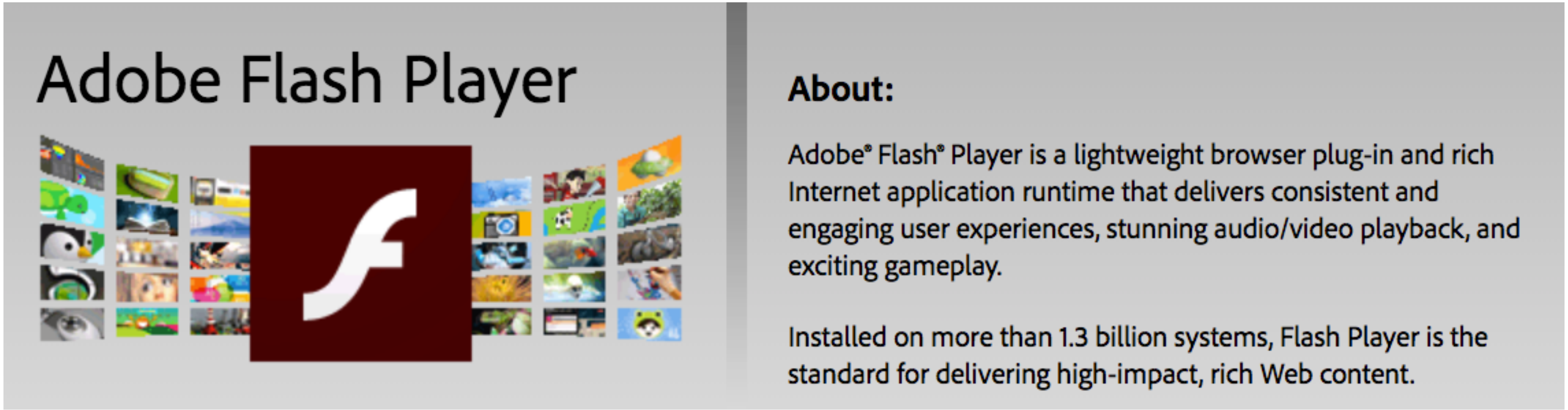
move presentation responsibility
to the client



Evolution 4: augment browser with an ad-hoc engine to be able to execute a (proprietary) language

Plug in: Adobe Flash
(former Micromedia Flash)

FLV-Video-Media-Content played by the Adobe Flash OCX Plugin.



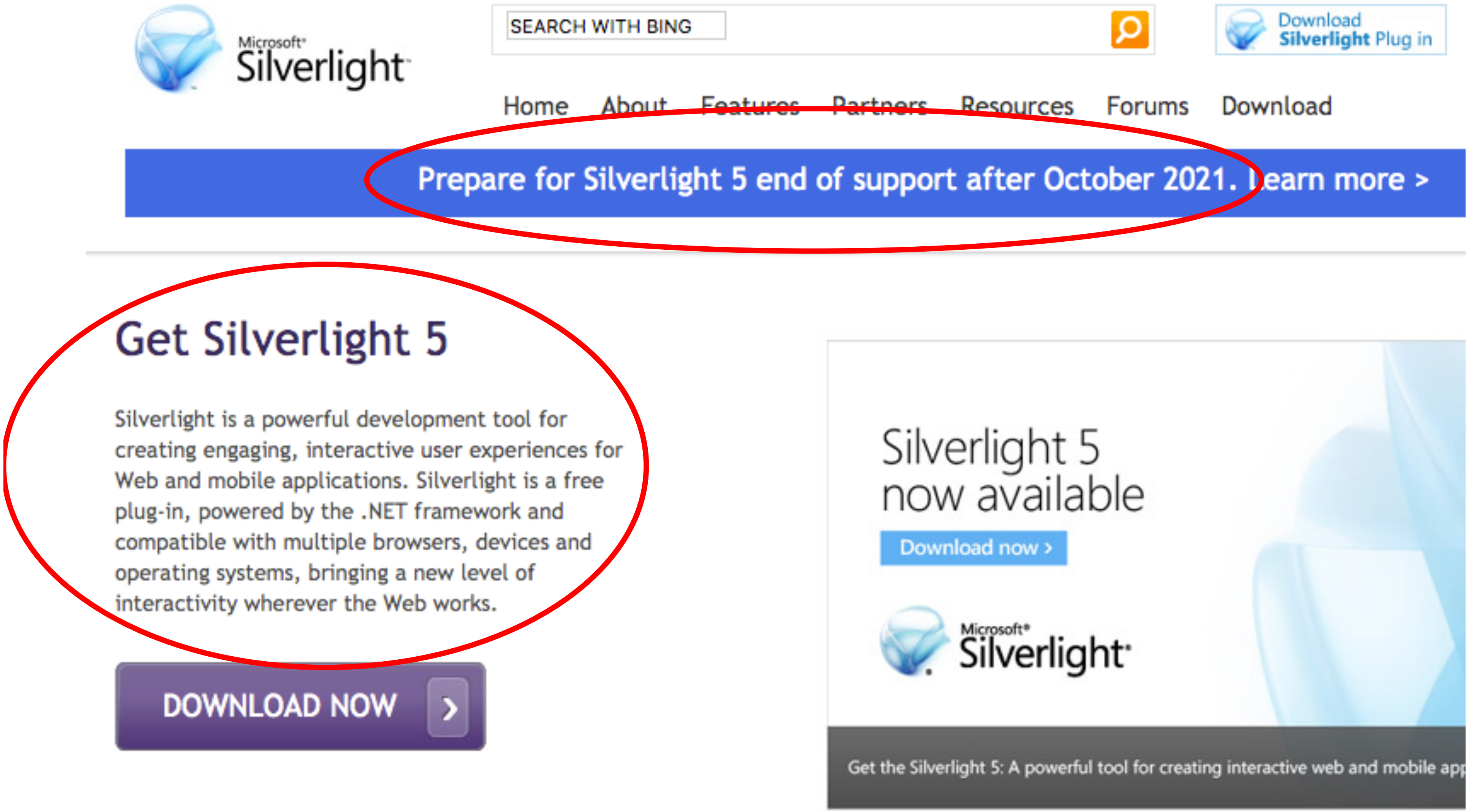
Adobe Flash Player

About:

Adobe® Flash® Player is a lightweight browser plug-in and rich Internet application runtime that delivers consistent and engaging user experiences, stunning audio/video playback, and exciting gameplay.

Installed on more than 1.3 billion systems, Flash Player is the standard for delivering high-impact, rich Web content.

Plug in: SIlverlight



Prepare for Silverlight 5 end of support after October 2021. Learn more >

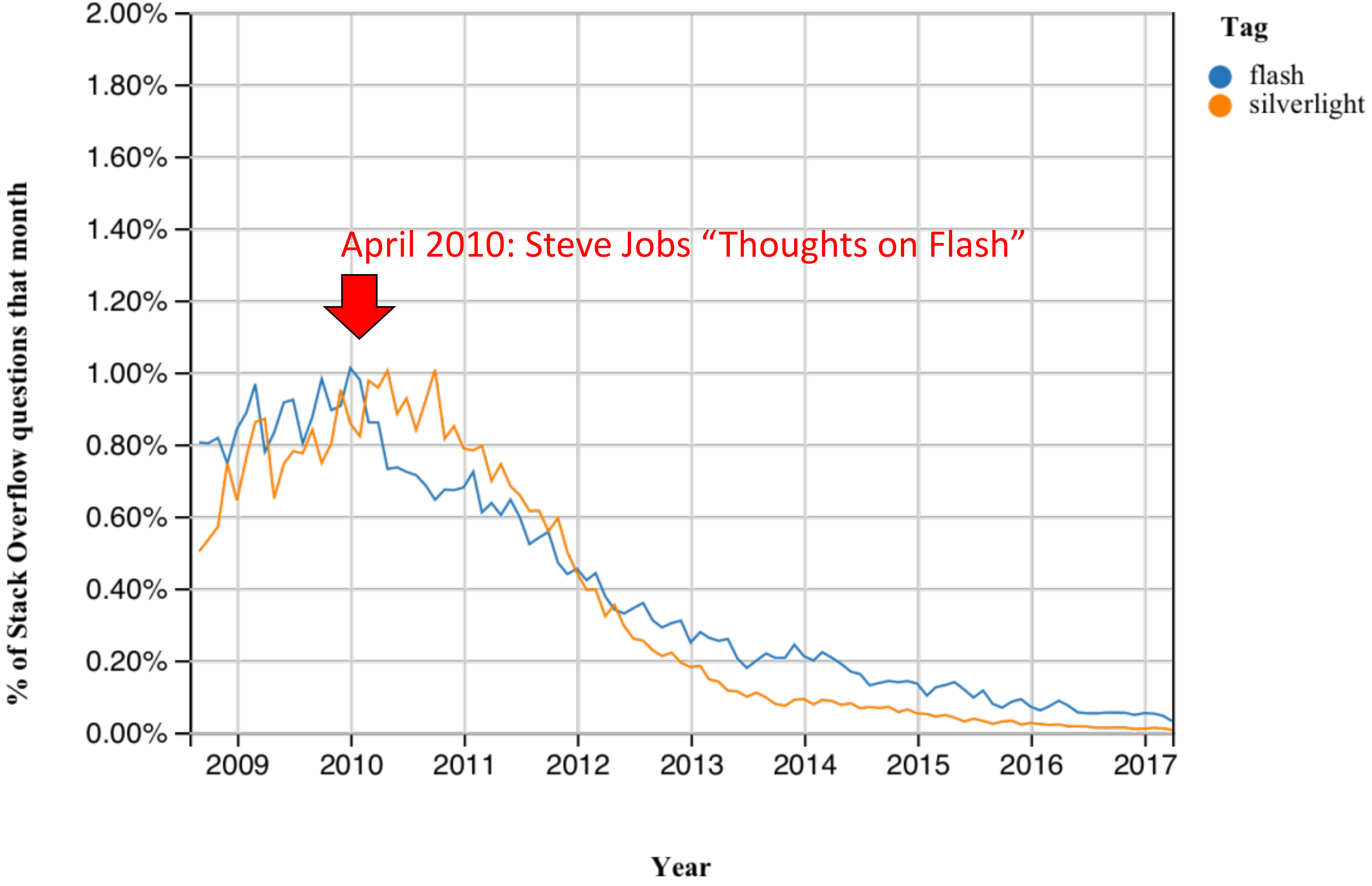
Get Silverlight 5

Silverlight is a powerful development tool for creating engaging, interactive user experiences for Web and mobile applications. Silverlight is a free plug-in, powered by the .NET framework and compatible with multiple browsers, devices and operating systems, bringing a new level of interactivity wherever the Web works.

Silverlight 5 now available

Get the Silverlight 5: A powerful tool for creating interactive web and mobile app

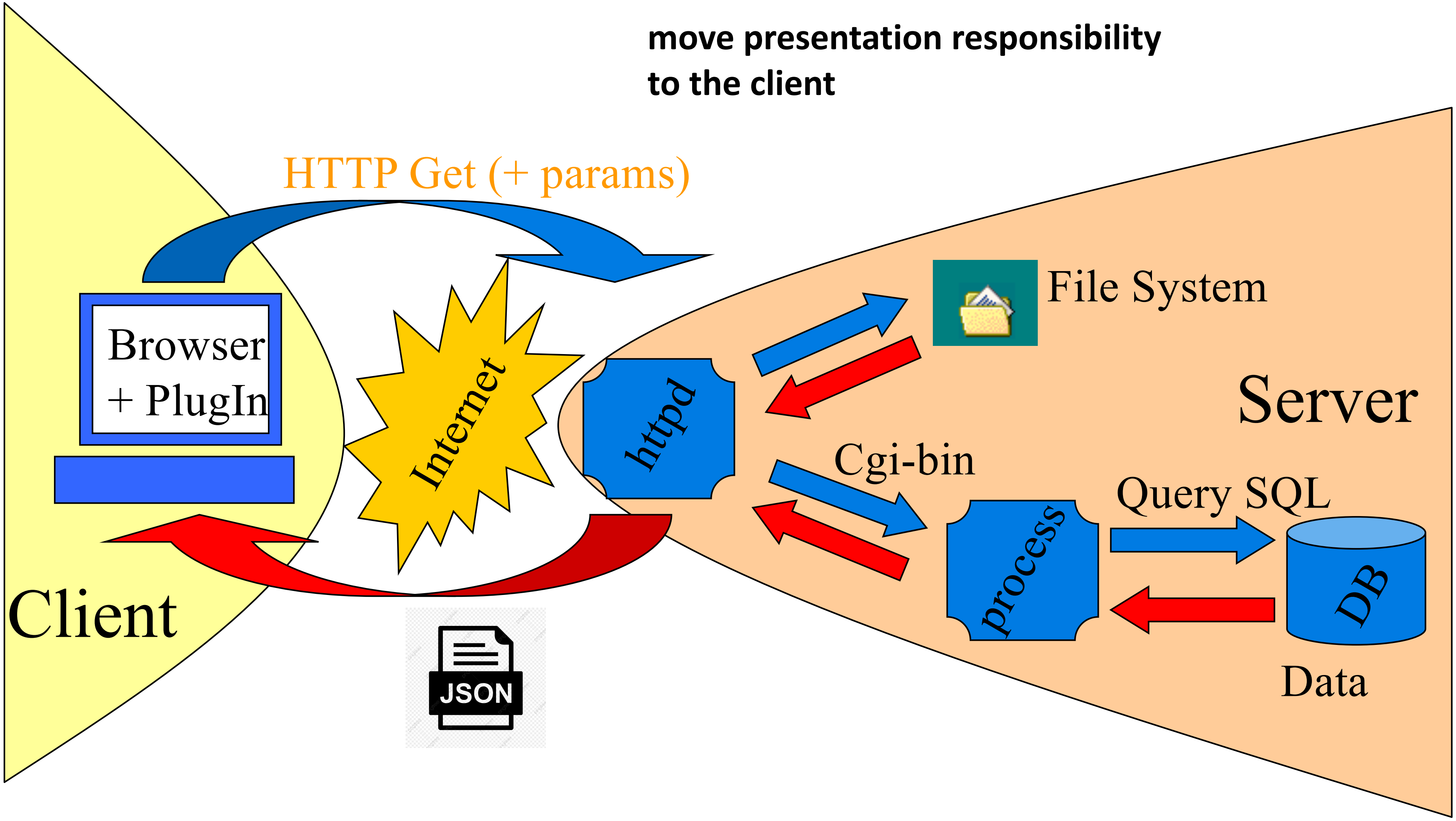
Flash and Silverlight decline



April 2010: Steve Jobs “Thoughts on Flash”

The web architecture: Ajax and SPA

move presentation responsibility to the client



Evolution 5: deliver data to single page applications

Q

What is a Protocol ?

Protocol

High-Change in Bond Street, —ou— la Politeſse du Grande Monde.

- **Synonymous of Etiquette**

a code of behavior that delineates expectations for social behavior according to contemporary conventional norms within a society, social class, or group.

Communications protocol, a set of rules and regulations that determine how data is transmitted in telecommunications and computer networking

Protocol

- *A **protocol** defines:*

1. *the format and*
2. *the order of messages exchanged between two or more communicating entities, as well as*
3. *the actions taken on the transmission and/or receipt of a message or other event.*

Q

What is a Port?

Port

A port is an **endpoint of communication in an operating system**.

A **process** associates its input or output channels, via an Internet **socket**, with a transport protocol, a port number, and an IP address.

socket: {protocol, local address, local port,
remote address, remote port}
This process is known as binding.

HTTP on port 80
- HTTP with SSL (HTTPS) on port 443
- FTP on port 21
- SMTP on port 25
- POP on port 110
- SSH on port 22

Mistranslated into Italian as "Porta" (door

PID	PORT	IP	Protocol
84	21	193.205.196.130	FTP
78	80	193.205.196.130	HTTP
321	8080	193.205.196.130	HTTP
541	25	193.205.196.130	SMTP

Q

What are URI, URL, URN ?

URI, URL, URN

- A **web resource**, or simply **resource**, is any identifiable thing, whether digital, physical, or abstract.

- A **Uniform Resource Identifier (URI)** is a compact sequence of characters that identifies an abstract or physical resource. (RFC 3986)
- A **Uniform Resource Locator (URL**) refers to the subset of URI that identify resources via a ***representation of their primary access mechanism*** (e.g., their network "location")
- **Uniform Resource Name (URN)** refers to the subset of URI that are required to **remain globally unique and persistent even when the resource ceases to exist** or becomes unavailable.

  It is intended to serve as persistent, location-independent, resource identifier (RFC 2141).

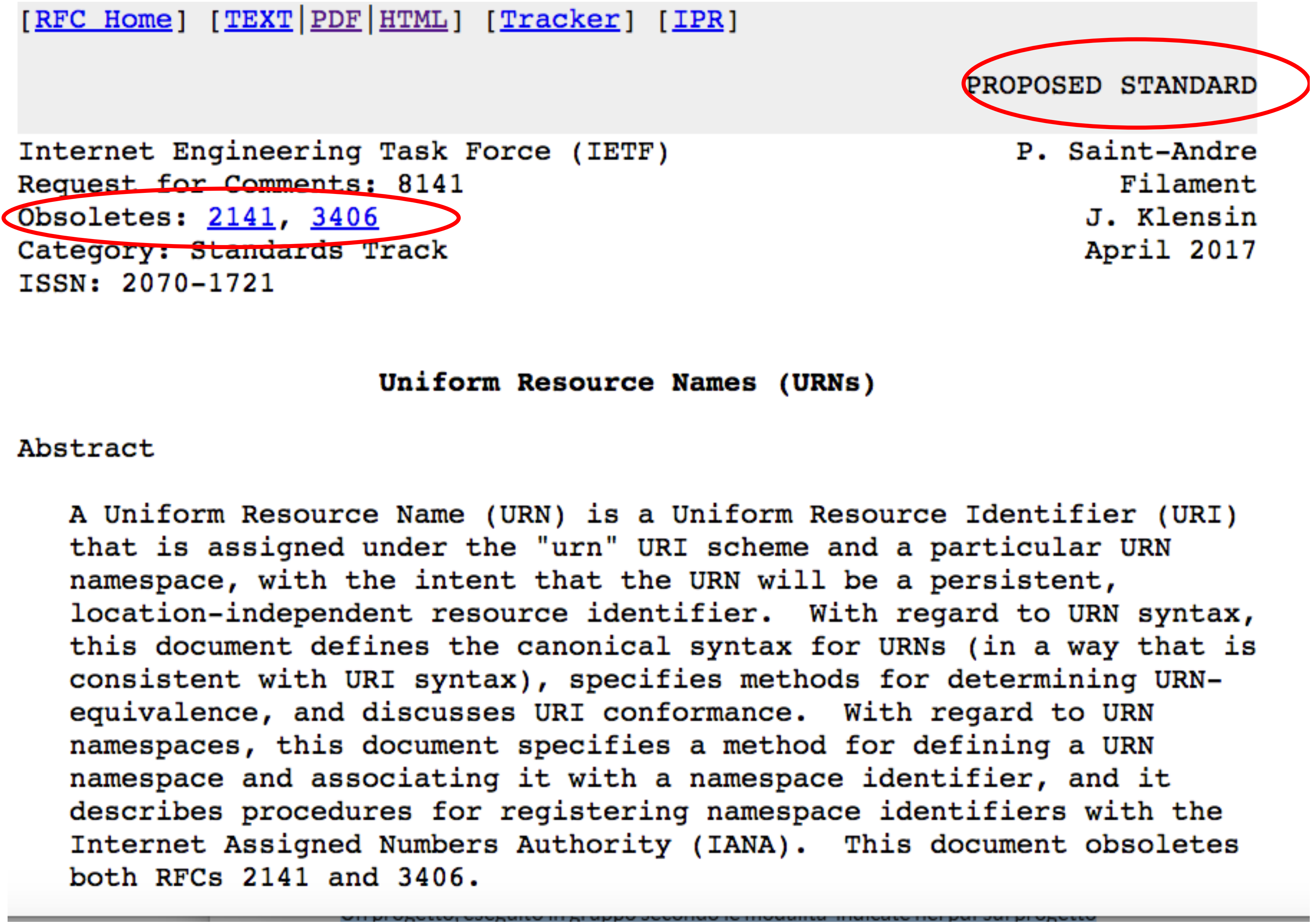
RFC (Request for comment)

Memos in the RFC document series contain technical and organizational notes about the Internet.

RFCs cover many aspects of computer networking, including protocols, procedures, programs, and concepts, as well as meeting notes, opinions, and sometimes humour.

https://ietf.org/standards/rfcs/

RFC



[RFC Home] [TEXT|PDF|HTML] [Tracker] [IPR]

PROPOSED STANDARD

Internet Engineering Task Force (IETF) P. Saint-Andre
Request for Comments: 8141 Filament
Obsoletes: 2141, 3406 J. Klensin
Category: Standards Track April 2017
ISSN: 2070-1721

Uniform Resource Names (URNs)

Abstract

A Uniform Resource Name (URN) is a Uniform Resource Identifier (URI) that is assigned under the "urn" URI scheme and a particular URN namespace, with the intent that the URN will be a persistent, location-independent resource identifier. With regard to URN syntax, this document defines the canonical syntax for URNs (in a way that is consistent with URI syntax), specifies methods for determining URN-equivalence, and discusses URI conformance. With regard to URN namespaces, this document specifies a method for defining a URN namespace and associating it with a namespace identifier, and it describes procedures for registering namespace identifiers with the Internet Assigned Numbers Authority (IANA). This document obsoletes both RFCs 2141 and 3406.

URI, URL, URN

- Both URL and URN are URI.

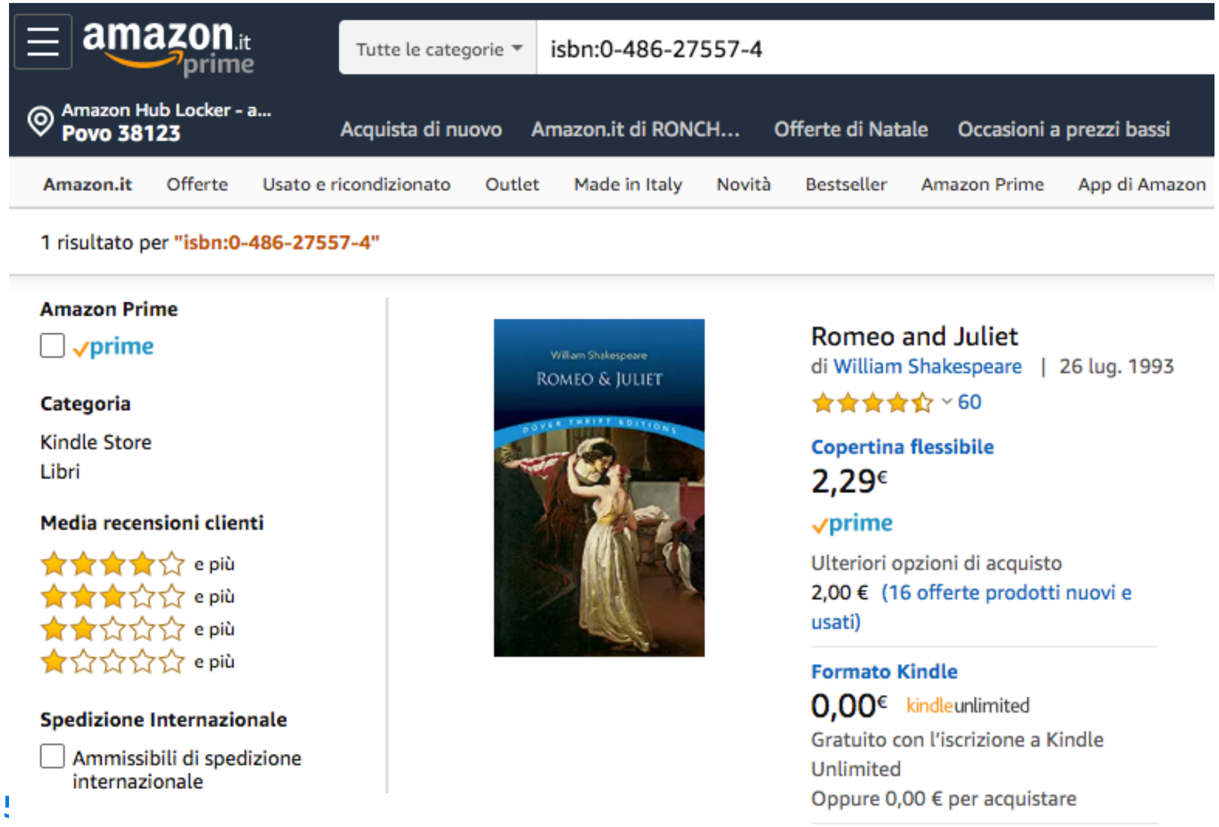
- A URN identifies a resource
- A URL provides a method for finding it.

- A URN may be compared to a person's name,
- A URL may be compared to their street address.

- A URN can be associated to many URLs

URN + URL example

- The ISBN system (namespace) uniquely identifies books.
- urn:isbn:0-486-27557-4 cites unambiguously a *specific edition* of Shakespeare’s play Romeo and Juliet.
- A typical URL for this book might look like the file path file:///home/username/RomeoAndJuliet.pdf



isbn:0-486-27557-4

1 risultato per "isbn:0-486-27557-4"

Amazon Prime

□ ✓prime

Categoria

Kindle Store
Libri

Media recensioni clienti

★★★★☆ e più
★★★☆☆ e più
★★☆☆☆ e più
★☆☆☆☆ e più

Spedizione Internazionale

□ Ammissibili di spedizione internazionale

Romeo and Juliet
di William Shakespeare | 26 lug. 1993
★★★★☆ 60

Copertina flessibile
2,29€
✓prime
Ulteriori opzioni di acquisto
2,00 € (16 offerte prodotti nuovi e usati)

Formato Kindle
0,00€ kindleunlimited
Gratuito con l'iscrizione a Kindle Unlimited
Oppure 0,00 € per acquistare

URI Schemes

http:

https:

ftp:

mailto:<address>[?<header1>=<value1>[&<header2>=<value2>]]

geo:<lat>,<lon>

fax:<phone number>

file:[//host]/path

bitcoin:<address>[?[amount=<size>]…

skype:<username|phonenumber>…

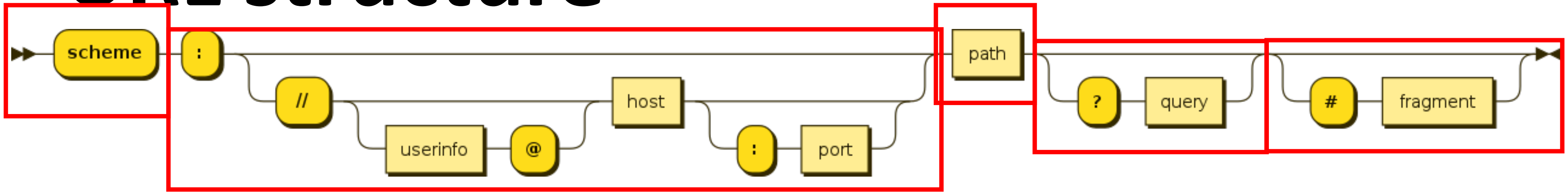
https://en.wikipedia.org/wiki/List_of_URI_schemes

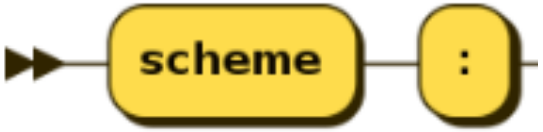
URI structure

URI = scheme:[//authority]path[?query][#fragment]

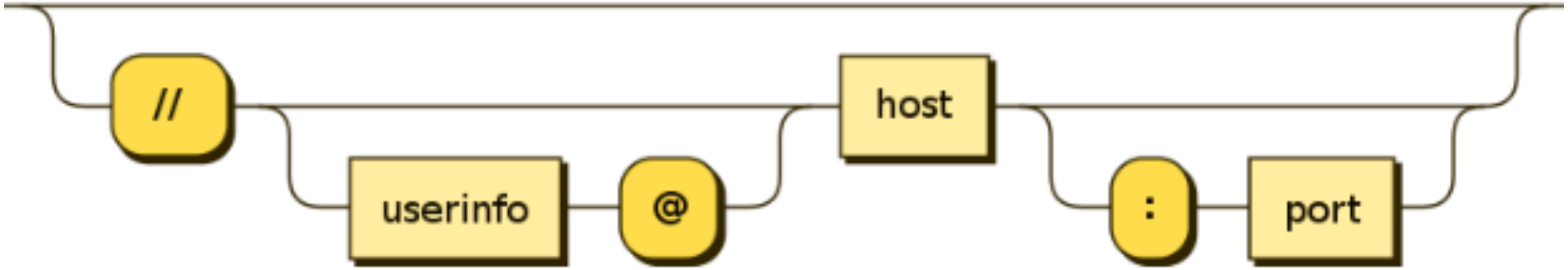
authority = [userinfo@]host[:port]

URL structure





https://en.wikipedia.org/wiki/URL

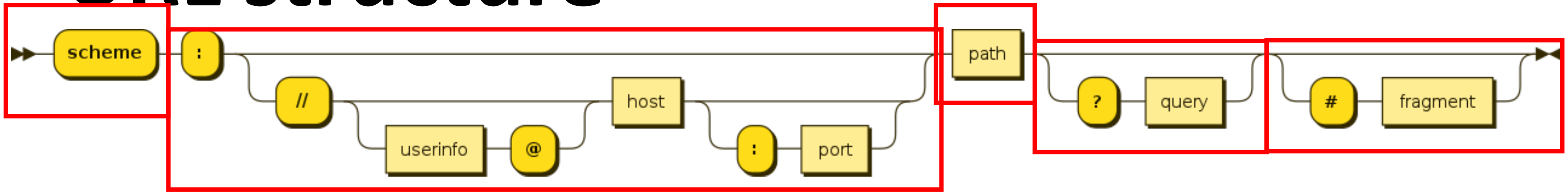


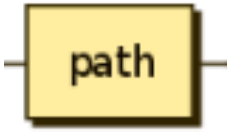
https://en.wikipedia.org/wiki/URL

Watch for phishing!

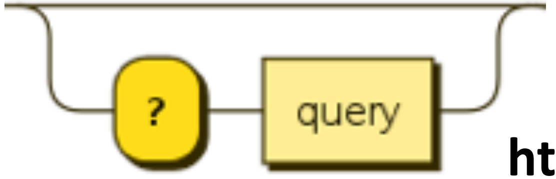
http://www.sitosicuro.it|search=hello@www.phishing.com/

URL structure

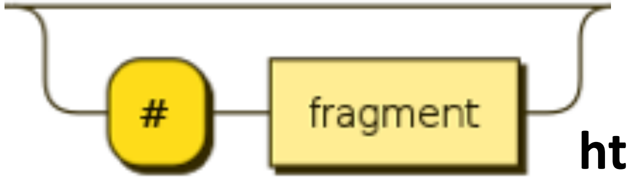




https://en.wikipedia.org/wiki/URL



https://en.wikipedia.org/w/index.php?title=URL&action=edit



https://en.wikipedia.org/wiki/URL#Protocol-relative_URLs

Q

What is MIME Type ?

MIME type

URIs give us info about name and location, but what about their content type?

File extensions are bad…

(see Unix magic numbers

https://www.geeksforgeeks.org/working-with-magic-numbers-in-linux/)

Metadata?

MIME type

Multipurpose Internet Mail Extensions (RFC 2045,2046)

MEDIA_TYPE/SUBTYPE

text -> text/plain, text/html, text/richtext …

image -> image/jpeg, image/png, image/svg+xml…

audio -> audio/basic, audio/ogg, audio/x-wav…

video -> video/mp4, video/ogg…

application -> application/x-apple-diskimage… …

multipart

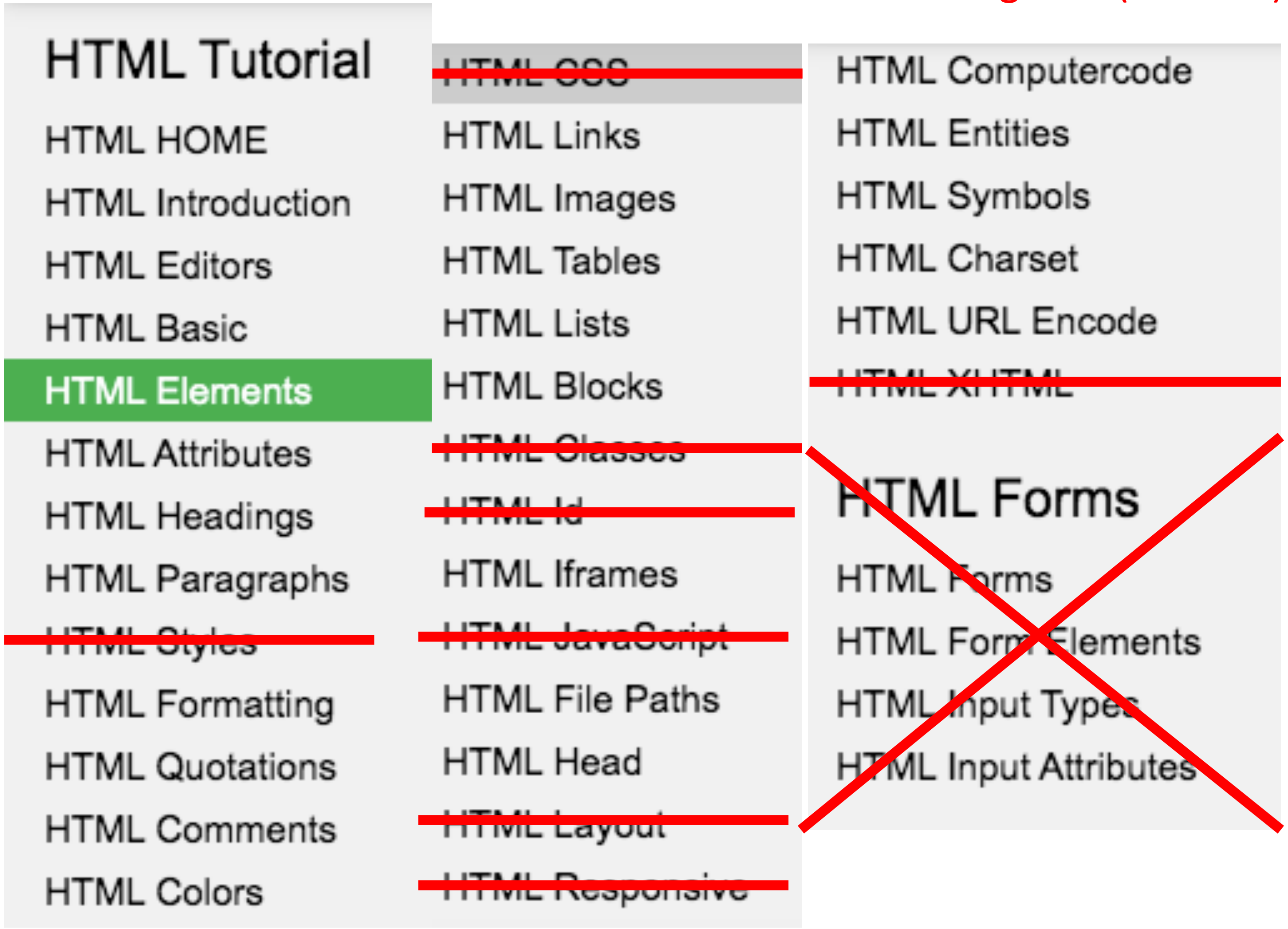
See https://www.freeformatter.com/mime-types-list.html

Q

What am I supposed to already know about HTML ?

Basic HTML

Ignore (for now) references to “styles”



HTML Tutorial

- HTML HOME
- HTML Introduction
- HTML Editors
- HTML Basic
- HTML Elements
- HTML Attributes
- HTML Headings
- HTML Paragraphs
- ~~HTML Styles~~
- HTML Formatting
- HTML Quotations
- HTML Comments
- HTML Colors
- ~~HTML CSS~~
- HTML Links
- HTML Images
- HTML Tables
- HTML Lists
- HTML Blocks
- ~~HTML Classes~~
- ~~HTML Id~~
- HTML Iframes
- ~~HTML JavaScript~~
- HTML File Paths
- HTML Head
- ~~HTML Layout~~
- ~~HTML Responsive~~
- HTML Computercode
- HTML Entities
- HTML Symbols
- HTML Charset
- HTML URL Encode
- ~~HTML XHTML~~

~~HTML Forms~~

- ~~HTML Forms~~
- ~~HTML Form Elements~~
- ~~HTML Input Types~~
- ~~HTML Input Attributes~~

https://www.w3schools.com/html/default.asp

Q

What are today's academic challenges in the web world ?

Open issues at today's conferences

Making the Web Human-Centric? New Directions in the Web and AI
12th International ACM Web Science Conference (July 2020)

Fairness, inclusion and diversity of Web and AI:

the construction of online identities, representation on the Web, access to the Web and technology, making 'smart' fair; Web culture and Web values.

Futures:

- the impact of the Web and new technologies on future society and social transformation,
- work futures and the data economy,
- health futures,
- political futures

Open issues: Safety, Security & Trust:

- Safeguarding and governance of the Web and/or AI;
- Anonymity, security and trust for Web access;
- False news;
- Data for the public good;
- Crime on the Web;
- Ethical challenges of technologies, data, algorithms, platforms, and people.

Open issues: Web & Contemporary Society

- Arts, culture and humanities on the Web;
- Web economics,
- Social entrepreneurship and innovation; health and online;
- Knowledge, education, and scholarship on/through the Web;
- Politics & political activism.

Open issues: Techno-social Web

- Modeling Web data, users
- Detecting, preventing and predicting anomalies in Web data (e.g., fake content, spam, algorithmic and data);
- Analysis and modelling of human vs. (e.g., bots) and their influence on the structure of the Web and responding ;
- Social machines, crowd computing, collective intelligence, and collaborative production on the Web.